DOCUMENT RESUME

ED 084 007 PS 005 697

AUTHOR	Freedle, Roy; Hall, William S.
TITLE	Effects of Prenominal Adjective Ordering on Children's Latencies and Errors in an Immediate Sentence Recall Task.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
REPORT NO	ETS-RB-73-19
PUB DATE	Mar 73
NOTE	29p.; An earlier version of this paper was presented at the Annual Convention of the American Psychological Association (80th, Honolulu, Hawaii, Sept. 2-8, 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Adjectives; Age Differences; Cognitive Development; *Early Childhood; *Imitation; *Language Development; Memory; Nominals; *Recall (Psychological); *Sentence Structure

ABSTRACT

A total of 34 children, ages 2 and a half to 6, were presented with sentences for imitation that either violated or honored a prenominal adjective ordering rule, which requires that size adjectives must precede color adjectives. Two response measures were evaluated in terms of these sentence types: latency to begin a sentence imitation and recall errors. For both the older and younger subjects, latencies following adjective order violations were significantly longer than following correct adjective order. This indicated the existence of a perceptual strategy in noun phrase segmentation which occurs at the time the sentence is comprehended. The recall error measure indicated that a different strategy is reflected in the output phase of a sentence imitation task: this strategy was called a "shift-to-grammatical-output." Older subjects were found to employ this latter strategy, whereas the younger subjects did not employ it. These results were interpreted in terms of Bever's developmental theory of prenominal adjective ordering acquisition, the empirical work of Martin and Molfese, and a more general developmental theory suggested by the work of Danks, Glucksberg, and Schwenk. (Author/SET)

FILMED FROM BEST AVAILABLE COPY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RB-73-19

ED 084007

RESEARCH BULLETIN

EFFECTS OF PRENOMINAL ADJECTIVE ORDERING ON CHILDREN'S LATENCIES AND ERRORS IN AN IMMEDIATE SENTENCE RECALL TASK

Roy Freedle
Educational Testing Service

and

William S. Hall
Princeton University

An earlier version of this paper was presented at the Annual Meeting of the American Psychological Association, Honolulu, Hawaii, September 1972.

PS 006697

Educational Testing Service
Princeton, New Jersey
March 1973

Effects of Prenominal Adjective Ordering on Children's Latencies and Errors in an Immediate Sentence Recall Task

Roy Freedle
Educational Testing Service
Princeton, N. J. 08540

and

William S. Hall
Princeton University
Princeton, N. J. 08540

Abstract

Thirty-four children ranging in age from 2,6 to 6,1 (median age 4,3) were presented with sentences for imitation which either violated or honored a prenominal adjective ordering rule which requires that size adjectives must precede color adjectives. Two response measures were evaluated in terms of these sentence types: latency to begin a sentence imitation and recall errors. For both the older and younger subjects, latencies following adjective order violations were significantly longer than following correct adjective order. This was argued to provide evidence for the existence of a perceptual strategy in noun phrase segmentation which occurs at the time the sentence is comprehended. The recall error measure indicated that a different strategy is reflected in the output phase of a sentence imitation task: this strategy was called a shift-to-grammatical-output. Older subjects were found to employ this latter strategy whereas the younger subjects did not employ it. These results were interpreted both in terms of Bever's developmental theory of prenominal adjective ordering acquisition and the empirical work of Martin and Molfese. Finally, these results were related to a more general developmental theory suggested by a pragmatic communication interpretation of adjective ordering due to the work of Danks, Glucksberg, and Schwenk.

Effects of Prenominal Adjective Ordering on Children's Latencies and Errors in an Immediate Sentence Recall Task[1,2]

Roy Freedle and William S. Hall

Educational Testing Service Princeton University

Princeton, N. J. 08540 Princeton, N. J. 08540

The purposes of this paper were twofold. The first of these was to provide empirical evidence regarding the existence (or absence) of a perceptual processing strategy which Bever (1970) has postulated occurs in sentence comprehension. In particular we shall be concerned with his strategy which facilitates noun phrase parsing at the time of sentence input. The second was to present a discussion of issues regarding the developmental course of this strategy, its possible age of acquisition, as well as its possible interaction with other behavioral strategies which occur at the time of overt recall of sentences, given that they have already been comprehended.

The paper begins by summarizing Bever's (1970) developmental theory of perceptual strategies for noun phrase parsing. Then empirical counter-evidence from the work of Martin and Molfese (1972) is recounted. In both instances prenominal adjective ordering was studied so as to provide data regarding the presence of a perceptual strategy for such segmentation. Following a brief critique of these studies, a possible resolution of some of the arguments stemming from the above papers is considered and new data are presented to test aspects of these conjectures. The paper concludes with a discussion of a more general developmental theory of prenominal adjective ordering strategies by implicating the pragmatic communication theory of Danks and Glucksberg (1971) and Danks and Schwenk (1972).

Bever's hypothesis. Bever (1970) maintains that in a given series of prenominal adjectives, the more nounlike adjectives are ordered so as to be closer to the head noun which they modify. For example, consider the fact that in English, adjectives of size (large, big, etc.) generally precede adjectives of color (red, blue, etc.). One can provide the following kinds of syntactic tests to illustrate that adjectives of color are more nounlike than adjectives of size. We cannot say Large is my preferred size* (where * indicates an ungrammatical sentence) but we can say Blue is my preferred color. Also we cannot say I dislike big the most*, whereas we can say I dislike red the most. In both pairs of examples we see that color words can function as nouns and so are more nounlike than size words. Bever suggests that this constraint may have emerged in order to facilitate segmentation of noun phrases during the process of ongoing speech comprehension. To show that such orderings do indeed facilitate segmentation consider the following two sentences: The large white steel rod broke versus The steel large white rod broke*. In the second case one expects to hear either a verb after "steel" or a head noun which will terminate the noun phrase. When a less nounlike adjective such as "large" follows it, we no longer can easily isolate where the end of the noun phrase will occur. On the other hand, for sentences in which adjectives are arranged by increasing nounness, this difficulty does not arise--hence the conclusion that prenominal adjectives which are constrained by increasing nounness order facilitate the segmentation of noun phrases in the speech signal.

While adult speech perception appears to be facilitated by this constraint it is by no means clear whether children are capable of utilizing this constraint in their perception of sentences. Bever has speculated that one might

find in a sentence recall task (Bever, 1970, p. 326) that younger subjects of about 2-3 1/2 years may not have internalized this constraint. If that is true he reasoned they should show no differences in their ability to recall immediately stimulus sentences which satisfy this constraint in contrast to sentences which violate this constraint. He further speculated that older subjects between about 3 1/2-5 years, who presumably have already acquired this rule, should experience greater difficulty in recalling sentences which violate this rule in comparison with sentences which satisfy it. In addition, Bever reasoned that the basis for this order of acquisition of the rule over age stems from the argument that one must first gain competence in comprehending or perceiving the occurrence of a regularity (such as the increasing nounness regularity in prenominal adjectives) before one can demonstrate competence in observing this rule in one's spontaneous sentence productions. This developmental argument for language has also been advanced by Fraser, Bellugi, and Brown (1963) who presented data suggesting that competence in imitation tasks precedes competence in comprehension tasks with production competence lagging behind comprehension and imitation abilities. Bever's argument differs from Fraser, Bellugi, and Brown's inasmuch as it allows for the possibility that comprehension strategies (such as is exemplified by the prenominal adjective strategy) and production preferences may all be implicated within the paradigm used for a sentence imitation task. Thus Bever has suggested that when a comprehension strategy has been violated one can detect this by examining the number of errors that occur in overtly imitating a sentence; also he has suggested that within this same sentence imitation task those error patterns which occur in the subjects' overt recalls of the sentences implicate production preferences.

While Bever reported some preliminary analyses of data which seemed to bear out his hypotheses, Martin and Molfese (1972) have recently reported results which appear to contradict his assertions. Let us examine this further.

The Martin and Molfese studies. The first study reported by Martin and Molfese[3] attempted to "replicate" the study mentioned by Bever (1970, p. 326). Two groups of children of approximately the same age range as indicated above were given a sentence recall task which used grammatical as well as ungrammatical prenominal adjective ordering to evaluate Bever's hypotheses. They failed to find evidence for differential knowledge of prenominal adjective ordering in examining the recall errors for these two types of sentences. They criticize Bever by suggesting that errors of recall need not reveal evidence for the existence or nonexistence of a perceptual strategy; instead they suggest that errors of recall may only reflect other types of behavioral strategies which do not necessarily have to implicate those strategies which occur during the process of sentence comprehension itself. To underscore their point, they carried out a second study which showed that the same children who were used in their first study were capable of revealing strong adjective order preferences in their sentence productions. This was obtained by telling the children that they had to give three names (adjectives) that described each of several houses. By studying the order with which the children gave these descriptive adjectives they found that size adjectives were preferred before color adjectives 82% of the time for both age groups. In addition, both groups showed a preference for giving size adjectives before cleanliness adjectives and color before cleanliness adjectives although the magnitude of the preference was somewhat less than the size-color preference.

A critique of the Martin and Molfese studies. While we concur with Martin and Molfese that error analysis at the time of sentence recall is probably not an appropriate measure of perceptual strategies which occur at the time of sentence comprehension, we feel that a number of procedural oversights somewhat attenuate the strength of several of their conclusions. To wit, in their first study Martin and Molfese did not actually replicate Bever's design and method. By introducing a physical toy on each trial and requiring that the child not only attempt to imitate the sentence on each trial but in addition imitate the action executed by the experimenter with the toy, they have introduced a potentially powerful distractor to the main purpose of their experiment--namely, to detect a differential effect in correctly recalling just the stimulus sentence. It is thus not surprising that they failed to get a similar effect to what Bever tentatively reported occurs in recall errors of the ungrammatical versus grammatical prenominal adjective orderings. More critically, though, is the fact that by introducing just a single object on each trial (a toy), the use of any qualifying adjectives (be they in correct or incorrect order) would seem quite superfluous. It has been argued (see Olson, 1970, 1972) that qualifying adjectives are typically used to communicate critical information so as to select one object from several alternatives in a set of possibilities. With but one toy present it may be quite redundant to use any qualifying adjectives.[4]

The second study of Martin and Molfese is a clear demonstration that children of about the age of 3 and older already possess definite adjective ordering preferences in their sentence productions. However, Martin and Molfese have overgeneralized the implications of their two studies by rejecting sentence imitation tasks as inappropriate for studying adjective ordering rules. They have overlooked the possibility that other response measures may

PS 006697

be defined in sentence imitation tasks which are sensitive to perceptual strategies at the time of initial sentence comprehension. We postulate that the latency interval between the end of stimulus sentence input to the point in time when the subject begins to imitate the sentence overtly is precisely the type of response measure which is likely to reflect the presence of Bever's noun phrase segmentation strategy. If this premise is correct, then one way to determine whether children of the older age group find ungrammatical sentences more difficult to comprehend than grammatical ones (i.e., they find it more difficult to begin an imitation of The brown large dog growled* than to begin an imitation of The large brown dog growled) is to demonstrate that the ungrammatical sentences require significantly more time to begin an imitation than the grammatical ones. If the subject who knows the prenominal adjective rule has been given a sentence which violates this ordering, then it should take more time for him to untangle the correct interpretation of this sentence than it would had the sentence satisfied the constraint. Thus, even if this hypothetical subject correctly recalled the stimulus sentence in exactly the form it was given him (he made no errors in his overt recall protocol), we might still hope to detect a difference between these grammatical and ungrammatical sentences by measuring the latency intervals. If our hypothetical subject is not sensitive to this difference (doesn't know the adjective ordering rule), then there should be no significant difference in the processing times for the ungrammatical versus the grammatical stimulus sentences. Finally, by also analyzing the types of errors made at the time of overt sentence recall we may also shed light on the nature of the behavioral strategies that exist across the two age groups in their overt recalls--this second response measure would allow us to compare error patterns in the

absence of a potential physical distractor such as the toy(s) used in the Martin and Molfese study. As such, it would provide a closer approximation to Bever's method than that employed by Martin and Molfese.

Figure 1 summarizes the hypotheses concerning the flow of events on a typical trial in a sentence imitation task and the response measures which can be obtained at different points of the trial; the figure also indicates how the two measures are believed to be implicated in detecting the presence (or absence) of a noun phrase segmentation strategy versus detecting the presence of other behavioral strategies such as might occur in the overt output phase of the trial. For example, one such strategy at the time of

---

Insert Figure 1 about here

---

overt output may be called a shift-to-grammatical-output in which the subject feels compelled to convert ungrammatical stimulus sentences into grammatical ones. For the ungrammatical sentence types to be used in this study (which will be elaborated in detail below) this strategy can be realized in several ways: the ungrammatical sentence The brown big dog growled* can be converted into a grammatical sentence by deleting the color adjective (yielding The big dog growled), the size adjective (The brown dog growled), both adjectives (The dog growled), or by inverting the order of the two adjectives (The big brown dog growled).

Materials

Five types of sentences were used. These were as follows:

1. The dog growled. (N)
2. The brown dog growled. (C)
3. The large dog growled. (S)

4. The large brown dog growled. (SC)

5. The brown large dog growled. (CS*)

where N represents sentences with no adjectives, C represents sentences with just a color adjective, S just a size adjective, SC sentences with a size and color adjective in that order, and CS* sentences with a color and size adjective in ungrammatical order.

By way of a rationale for this choice of stimulus sentences, one of the pertinent analyses will allow us to test whether SC* sentences take significantly longer to process than SC sentences; but over and above this, we can inquire as to just what the relationship is among these sentence types. One way to think about this is to consider sentences of types N, C, and S to represent basic building blocks out of which the more complex sentences can be formed. For example, one might postulate that the behavioral data will favor some linear combination of N, S, and C type sentences in predicting the results for SC type sentences. That is, one can use N, S, and C sentences as predictor variables in a regression analysis with SC sentences as the criterion variable to be predicted. In addition, using the same three predictors (and interaction terms as well, to be described below), one can study the relationship of N, S, and C type sentences to CS* type sentences where the latter is made the criterion variable to be predicted. If younger subjects are insensitive to the ungrammatical character of CS* sentences (or also if they do not use Bever's perceptual strategy), then the linear combination of beta weights for N, S, and C predictors should be indistinguishable from the beta weights for these same predictors when SC sentences are to be predicted. If older subjects (according to Bever's hypothesis) already are sensitive to the prenominal ordering constraints on adjectives, then a significantly

different pattern of beta weights should occur when CS* sentences are being predicted from N, S, and C sentences than when SC sentences are being predicted.

Method

Twenty sentences were presented to each of 34 subjects with four exemplars of each of the five sentence types being presented to each subject. Five test forms were constructed with test for s being randomly assigned to each subject. To illustrate how a given sentence was distributed among the five forms, consider the five sentences just given above. The first sentence (The dog growled) was assigned to the first test form, the second sentence to the second form, etc. This assignment of sentences to forms was continued until each form had exactly four exemplars of each of the five sentence types. All the sentences were of the active-affirmative intransitive type with 10 inanimate and 10 animate count nouns used as subjects of the sentences.

Procedure

Thirty-four subjects from a nursery school in the Princeton area (15 females, 19 males) were tested; they were all white middle-class children with English as their native language. Their ages ranged from 2,6 to 6,1 with a median age of 4,3.

Testing took approximately 15 minutes for each subject. Three short sentences were given as a warm-up prior to beginning the study proper. If the subject failed to say all the words in these practice sentences, he was asked to try again following another repetition of the sentence. All sentences were read with as natural intonation as possible. The entire

session was tape-recorded. Instructions were as follows: "I am going to say some things and I want you to say the same things that I say." This was followed by the warm-up sentences which were preceded by the special instruction "Say, ..." During the experiment proper, none of the sentences were prefaced by "Say." There was a two-second intertrial pause. If a child failed to say all the key content words on his first attempt at recall, he was read the same sentence and asked to try again--this was done to discourage fragmentary responding.

Following a typed transcript of each child's session, two judges used stop watches and timed the interval from the end of the stimulus sentence to the beginning of the subject's utterance. The tapes were played back at one-fourth their original speed so as to increase the accuracy of the latency scores. Ninety percent of the two judges' scores were within .05 sec (real-time) of one another; the mean of these two scores were averaged and used as the final estimate for each sentence. Only first attempts at recall were analyzed (only three instances of failure to respond on first attempts occurred; these trials were not included in the subject's mean latency for the sentence types for which the omission occurred).

Results

Latency analyses. Five latency scores were obtained from each subject which represented the mean time to respond to the four exemplars for each of the five sentence types. If a particular subject did not attempt to imitate any sentence or a part of the sentence, this trial did not enter into the calculation of the mean latency score for the sentence type in which it occurred, since no latency score could be estimated for such a trial. (There

were extremely few instances where the subject didn't attempt some aspect of sentence imitation.) The mean latency to begin a sentence recall for the younger subjects (ranging from 2,6 to 4,3) was .633 sec, .648 sec, .640 sec, .673 sec, and .762 sec (real time) for sentence types N, S, C, SC, and CS*, respectively. For the same order, the older subjects' (ranging from 4,4 to 6,1 in age) mean latencies were: .510 sec, .531 sec, .539 sec, .681 sec, and .701 sec. Thus, while the younger subjects, as evidenced by their mean latencies, tended to be slower in beginning an imitation, they were not significantly slower as evidenced by a Mann-Whitney U test ($U = 103$, $p > .10$, two-tailed), which used as a score for each subject the sum of the mean latencies over all five sentence types. We shall see later, though, in a more detailed analysis that age differences do occur as a function of the particular sentence types.

The regression analyses[5] showed clear-cut differences between the grammatical SC sentences and the ungrammatical CS* sentences. The following format was used in detecting this difference. When SC was used as the criterion variable and with N, S, C, NxS, NxC, SxC, NxSxC, and Age used as predictor variables, Table 1 demonstrates that the overall F test of the significance level for relating the criterion to the predictors was highly significant [$F(8,25) = 4.08$, $p < .005$]. In contrast, with CS* as the criterion and with the same predictor variables, the overall F test is not significant [$F(8.25) = 2.00$, $p > .10$]. Furthermore a more detailed look at the magnitude

---

Insert Table 1 about here

---

of the t values for each beta weight (with SC as the criterion) shows that the Age variable (which scored each of the 34 subjects with a "1" if he or she

were in the younger half of the sample, otherwise he or she got a "0") as well as the S scores, and the interaction measure scores SxC and NxSxC are significantly related to the criterion ($p < .05$, two-tailed, for each comparison). But when we examine the t values for each beta weight with CS* as the criterion, we see that none of the betas are significant! This dramatic difference in the regression results indicates that both younger and older subjects detect the anomalous character of CS* sentences in comparison with SC sentences; another way to view this result is to emphasize that all of the predictor variables (except Age) reflect latency scores for responding to grammatical stimulus sentences; when the criterion is also grammatical (by virtue of the adjective ordering constraint), then a significant F test relates the criterion to the predictors. But when the criterion is ungrammatical (by virtue of violating the adjective ordering constraint), then a nonsignificant F is obtained.

The other aspect to these regression results reflects upon our earlier conjecture (in the introduction) that the latencies for grammatical SC sentences may be represented by some simple linear combination of the latency scores obtained for the "basic building block" sentences of types N, C, and S. The regression equation for SC as criterion shows that a more complex function relates SC to its "basic building block" sentence types; in particular, it indicates that primarily the latencies for S type sentences in combination with such cross-product terms as SxC plus NxSxC are the significant contributors which best "reconstruct" the observed latencies for SC sentences.

Also, the regression analyses indicate that the two Age groups differ somewhat in the relative contribution which the predictor variables make in predicting the grammatical criterion SC sentences. Subanalyses suggested that

while the same predictor variables emerged as the significant predictors of SC within each age group, the relative contribution of the S predictor is greater for the older subjects while a smaller weight is contributed by the younger subjects. Although of interest in itself, we shall not dwell on it since it does not bear that directly upon the major issues concerning sensitivity to adjective ordering with respect to age.

The regression analyses have informed us that the pattern of latencies is significantly different depending upon whether the criterion variable satisfied the adjective ordering constraint or violated it. But this does not tell us whether CS* take significantly longer to process than SC sentences. A t-test indicated that over all subjects, CS* takes significantly longer to respond to than SC sentences ($t = 1.72$, $p < .05$, one-tailed). A further examination of the mean latencies indicated that both the young and older subjects contributed equally to this significant effect--thus it appears that both young and old subjects in the age range 2,6 to 6,1 are sensitive to the ungrammatical character of CS* sentences. By our hypotheses described earlier in this paper, the significantly longer latency to respond to CS* suggests that these subjects must be using Bever's noun phrase segmentation strategy. To recapitulate the argument: if a subject uses the perceptual strategy for identifying the end of noun phrases in order to facilitate sentence comprehension, and if a particular stimulus sentence violates this strategy, then it should take somewhat longer to comprehend this sentence inasmuch as he will have to do extra work in parsing the noun phrase part of the sentence. Our latency measure appears then to provide evidence that our younger (and older) subjects do indeed employ this segmentation strategy since otherwise it would be difficult to account for

both the significantly longer latencies for CS* sentences and the significantly different pattern of beta weights which occur when CS* versus SC sentences are used as the criterion variable in the regression analyses.

Error analyses. Congruent with the suggestion of Martin and Molfese that error analyses probably reflect behavioral strategies other than Bever's segmentation strategy for noun phrase parsing, an error analysis of just the noun phrase portion of the overtly recalled sentences was undertaken. In particular we evaluated the possibility that one behavioral strategy that may exist in the recall phase was a shift-to-grammatical-structure strategy in which a subject would tend to convert ungrammatical sentences into grammatical ones. Such an analysis would be of interest for several reasons. First, we have already seen that both younger and older subjects appear to employ Bever's noun phrase parsing strategy at the time the sentence is comprehended; it is relevant to inquire then whether both groups, just one, or neither show any evidence of sensitivity to CS* versus SC sentence differences when just their recall error patterns are examined. While other error analyses could have been undertaken, the shift-to-grammatical-structure one is especially relevant to studying this connection of recall protocols vis-à-vis Bever's perceptual strategy.

A sentence was counted as a positive instance of the shift-to-grammatical-output if the subject converted a sentence such as "The brown big dog growled*" into "...brown dog ...", "...big dog ...","... dog growled" or "...big brown dog...." That is, we did not require that the article be present or correct nor did we require that the verb be present or correct in scoring for the presence of this strategy. As a control condition, a similar scoring procedure was applied to the grammatical SC sentences to determine how often these deletion patterns (or inversion patterns as in converting "brown big dog" into

"big brown dog") occur for grammatical sentences with the same elements present. A within-subjects comparison was made by tallying the number of such shifts out of a maximum of four for CS* type sentences and comparing this with the number of such spontaneously occurring deletion patterns (or inversions) that occurred for CS type sentences. For the younger group of subjects a nonparametric sign test failed to show any evidence for the presence of this strategy ($p = .377$, one-tailed sign test). However, a similar comparison made for the older subjects revealed a significant difference in the use of this shift-to-grammatical-output strategy ($p = .006$, one-tailed sign test), with 7.3% of the scorable trials showing the "shift" for the grammatical SC sentences versus 23.5% showing the shift for the ungrammatical CS* sentences. The comparable percentages for the younger group were 22.4% and 26.5%, respectively. (Incidentally, in both groups virtually all of the scorable instances of this strategy resulted from the deletion of a single adjective--there were only two instances of reversing the two adjectives to make a grammatical utterance.) This result suggests that while both groups appear to employ the perceptual strategy for noun phrase segmentation at the time of sentence comprehension, they can employ quite different strategies when it comes time to recall the sentences overtly. Thus this result is consistent with Martin and Molfese's claim that recall error patterns need not implicate the presence or absence of the perceptual strategy which is presumed to occur at the time of sentence comprehension. Instead, a different response measure must be used to detect the perceptual strategy--latencies.

Discussion

In light of the above analyses of latencies and error patterns that occur in a sentence imitation task designed to study the existence or absence of a perceptual strategy for facilitating noun phrase segmentation during comprehension, we have found that both our younger and older groups of subjects appear to utilize this strategy at the time of sentence comprehension but employ different response strategies at the time of overt recall of the sentences. These results in combination with Martin and Molfese's production preference study for adjective ordering indicates that subjects of about age 2,6 already possess clear-cut preferences in their spontaneous sentence productions as well as perceptual strategies in comprehending sentences. These two results appear to pose a special problem for Bever's further conjectures regarding the necessity of learning these perceptual strategies as a developmentally prior period to the honoring of these grammatical constraints in one's spontaneous productions. (This last issue refers to the Fraser et al. (1963) conjecture that comprehension precedes production ability.) However, further thought indicates that neither the present study nor the two by Martin and Molfese actually contradict Bever's claim. The reason is that in order to demonstrate that perceptual strategies in comprehension are prior in development to production preferences, one must have data which indicate that for a given age group <u>just</u> the perceptual strategies are present in comprehension but do not reveal themselves in production, whereas older subjects reveal both the presence of the perceptual strategies and the production preferences.[6] The present study shows that for the age range studied (2,6 to 6,1) the subjects are already <u>too old</u> inasmuch as both groups appear already to employ the segmentation strategy for noun phrases;

this, in conjunction with Martin and Molfese's production study for approximately the same age groups, indicates that an appropriate test of Bever's claim about developmental sequencing of comprehension strategy being followed by production constraints will have to await the testing of even younger subjects than employed heretofore.

Further developmental issues. Up to this point we have purposely avoided mentioning various competing theories which purport to explain or account for observed adjective ordering constraints in prenominal adjectives. To test the particular issues which were at stake it was not necessary to summarize the diverse "rules" which attempt to account for the preferred orderings--it was sufficient to know that it was generally agreed that size precedes color adjectives and then to move forward from this to a consideration of whether different age groups have internalized this regularity. However, some recent studies by Danks and Glucksberg (1971) and Danks and Schwenk (1972) have opened up some new theoretical possibilities for discussing the developmental issue of comprehension and production competences. While the empirical work of Danks, Glucksberg, and Schwenk has concentrated on adult subjects, the regularities which they have uncovered can be examined from a developmental standpoint as we shall now attempt to make clear.

The work of Danks, Glucksberg and Schwenk has isolated a general pragmatic-communication rule which includes the particular rule for size-color ordering (among other orderings such as the size-cleanliness and color-cleanliness orderings, etc.) as the most frequently occurring special case of this more general pragmatic rule. In particular, they showed that by manipulating the set of alternatives from which a particular item was to be designated via prenominal adjectives, subjects preferred to place the critical

adjective first--wherein the adjective which was critical was a function of the set of alternatives in the choice set which varied in their attribute dimensions. By altering the set of alternatives so that sometimes the color attribute was critical, the subjects preferred placing this attribute first among the prenominal adjectives, provided that stress intonation was given this critical attribute. Hence the pragmatic communication context can override the preferred ordering that occurs in the most frequently encountered special case of this rule. While other interesting results were obtained by Danks, Glucksberg, and Schwenk, we shall focus upon the developmental implications of their work.

Since the work of Danks, Glucksberg, and Schwenk suggest that adults use this more general pragmatic rule, this raises the question as to whether young children also use this very general production "rule" or whether they only operate with the more restricted special cases of the rule--such as using the size-color ordering regardless of the situational context. Similarly, for comprehension tasks, one may inquire whether only the special case has been learned rather than the general one isolated by Danks, Glucksberg, and Schwenk.

One can identify four aspects to this general developmental sketch of the hypothesized transition from knowing the special case to knowing the contextual pragmatic rule: (A) in comprehension, the subject employs perceptual strategies which reflect only the special case of the rule; (B) in production preferences, the subject employs only the special case of the rule; (C) in comprehension, the subject shows sensitivity to the set of alternatives and hence shows knowledge of the general pragmatic communication rule; and (D) in production preferences, the subject reflects a sensitivity to the set of alternatives and hence is sensitive to the pragmatic context. Empirical studies may be fruitfully pursued by using a modification of the Martin and Molfese production task so as

to present a set of alternatives per trial from which a description is to be generated by the subject. This is to be contrasted with descriptions which result from trials wherein but a single object is presented. The results of such studies for both production preferences and comprehension (the comprehension tasks would be elaborations of a sentence imitation task which uses successive sentences to establish a pragmatic context effect) should tell us whether it is reasonable to assume that the stages of acquisition are such that A develops first, followed by B (thus far the ordering would be consistent with Bever's claim that comprehension competence precedes production competence for the special case of the pragmatic rule), followed by C, and finally by D. Other orderings of these four "stages" might be postulated, but the one suggested above (ABCD) seems the most likely one for the following reasons. One can identify an analogous process to the above ordering which has received considerable support in the empirical literature: that is the use of what is called a canonical ordering strategy (see Bever, 1970, for a review of this literature). The canonical strategy is apparently arrived at by overgeneralizing the pattern of active clauses which occur with high frequency in the language. Thus active sentences like Mary hit Bill which consists of a noun-verb-noun (NVN) pattern in the surface structure results in the assignment of the first noun to the subject of the sentence while the second noun is interpreted as the object. The canonical strategy says that any NVN sequence will be interpreted as subject-verb-object. This overgeneralization can lead to errors when applied indiscriminately. For example, passive sentences, which are of a lower frequency of occurrence in the language, also have a NVN sequence (e.g., Bill was hit by Mary). The correct assignment of deep structural relations for such passive

sentences is object-verb-subject. The application of the canonical strategy leads to the incorrect assignment of subject-verb-object to such sequences. Bever has indicated that older subjects gradually come to realize that not every NVN sequence should be interpreted as subject-verb-object. Thus with time the overgeneralized strategy gets replaced by a more differentiated context sensitive process. We can identify a similar process for the hypothesized transition from the special case of the prenominal adjective ordering rule to the pragmatic context-sensitive rule employed by the adults in the experiments of Danks, Glucksberg, and Schwenk. For the developmental order ABCD suggested above, the AB portion is seen to precede the CD portion. Since the AB part reflects the application of an overgeneralization of the most frequently occurring case of adjective ordering (which DGS has called the special case of the pragmatic rule), while the CD part reflects the application of a context-sensitive rule which is more differentiated than the special case, we see that a similar pattern is involved here as was uncovered for the canonical ordering strategy. First the young subjects overgeneralize by induction from the most frequently occurring pattern in the language, and later with older subjects this is followed by a more differentiated strategy which reflects a sensitivity to the context in which the sentence occurs.

References

Bever, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970. Pp. 279-362.

Danks, J. H., & Glucksberg, S. Psychological scaling of adjective orders. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 63-67.

Danks, J. H., & Schwenk, M. A. Prenominal adjective order and communication context. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 183-187.

Fraser, C., Bellugi, U., & Brown, R. Control of grammar in imitation, comprehension, and production. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 121-135.

Martin, J. E., & Molfese, D. L. Preferred adjective ordering in very young children. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 287-292.

Olson, D. R. Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 1970, 4, 257-273.

Olson, D. R. Language use for communicating, instructing, and thinking. In R. O. Freedle & J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: V. H. Winston, 1972. Pp. 139-167.

Footnotes

[1]This research was supported in part by the National Institute of Child Health and Human Development, Grant 5 P01 HD01762 to the Educational Testing Service and NIH Bio-Med Support Grant to Princeton University. An early version of this paper was presented at the annual American Psychological Association meetings held in Honolulu, Hawaii, September, 1972.

[2]The authors wish to express their appreciation to Margaret Nancy White and James Tittimore for helping with the experiment, data analyses, and tape transcriptions. Special thanks are due Susan Weiner for her critical comments on an earlier draft of this paper.

[3]Inquiries from Professor Martin, initiated after the completion of the research reported here, indicated that a failure in communication occurred between Martin and Molfese and Bever regarding the method used in the Bever-Epstein experiment. It is our understanding that physical toys were present in the main Bever-Epstein study; however, acting out with the toy was not required (Bever, personal communication). Hence Martin and Molfese's study was not an exact replication of the experiment reported by Bever (1970). Further communication with Bever indicated that additional data were obtained without toys present. The absence of a toy appeared to result primarily in more sentence deletions than when toys were present (Bever, personal communication).

[4]Actually, there appears to be a class of instances where just one object is overtly present but this evokes special comment expressed via the qualifying adjectives because the particular object departs from either an expected one (as in saying "Oh, look at the small piece of pie you gave me" given that one on that occasion expected to receive a large piece) or because

it departs from some idealized image of the object (as in saying "Look at the blue firetruck" where one adds the "qualifying" information about color because this particular instance of a firetruck departs from the idealized one across all occasions wherein every firetruck has been red). Thus a more inclusive theory of adjective use would require knowledge of what overt alternatives are present from which a particular one is to be designated, as well as a knowledge of what characteristics are expected to be selected, plus a knowledge of whether the overtly present alternative(s) is in line with the idealized range of values typically taken on by these objects in one's everyday experiences. Hence external givens plus internal knowledge and expectations must be combined into a general theory of message construction and word-object reference. To develop such a theory in detail in this paper would take us too far from the main goals; it will be developed elsewhere.

[5]A more efficient procedure for relating possible age effects to SC and CS* sentences than the argument given in the introduction was actually employed; rather than doing separate analyses for each age group, the combined sample of subjects was used with an age variable added so as to reflect possible differences due to age--this approach provided a more efficient assignment of beta weights to the various predictor variables as well as allowing for a test of significant age effects. When a significant age effect occurred in the present study, subanalyses with each age group analyzed separately indicated that the same predictor variables were significant in each age group but the magnitudes of respective beta weights were significantly different which led to a significant difference for the age predictor using the total sample of subjects.

[6]While the shift-to-grammatical output for the older subjects can be interpreted as evidence for a production preference, the absence of this output by the younger subjects cannot be said to imply the lack of a production preference, especially since Martin and Molfese's second study showed that young subjects already possess adjective ordering preferences in production. Hence this points up the fact that the output phrase of an imitation task need not coincide with the results obtained from a real production task.

Table 1

Regression Analyses

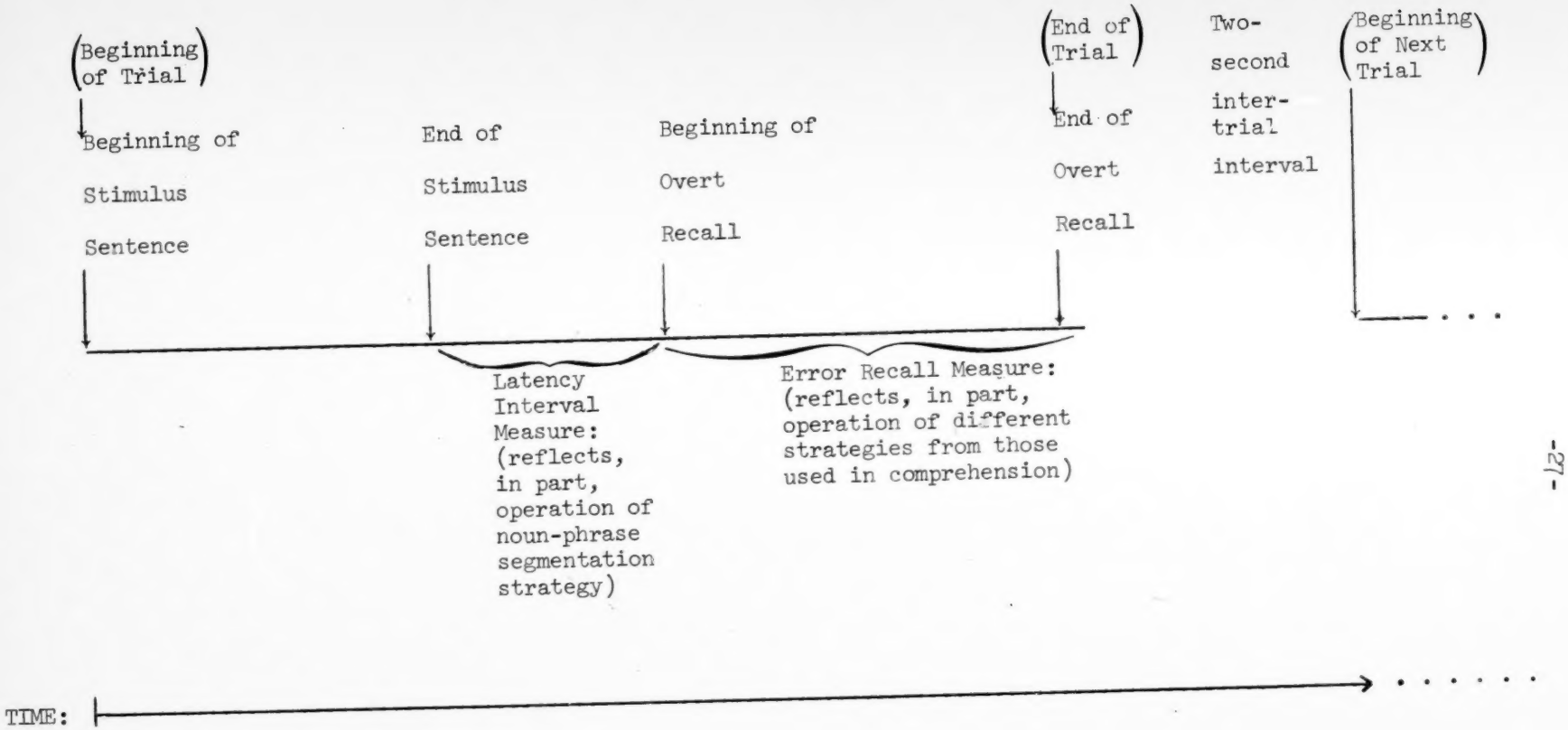
Criterion Variable	Beta Weights for Predictor Variables								Added Constant	F Ratio	p	Multiple R^2
	N	S	C	NxS	NxC	SxC	NxSxC	Age				
Grammatical SC sentences	1.8	2.1*	1.6	-1.9	-1.4	-2.2*	2.0*	-2.5**	-24.2	4.08	<.005	.566
Ungrammatical CS* sentences	0.1	0.8	0.2	-0.3	0.0	-0.6	0.4	-0.4	0.6	2.00	<.100	.390

*Indicates that beta weight is significant at the .05 level, two-tailed test.

**Indicates that beta weight is significant at the .02 level, two-tailed test.

Figure Caption

Fig. 1. Relationship between the latency measure, the error measure, and the flow of events on a typical trial.


(Beginning of Trial)
Beginning of Stimulus Sentence
End of Stimulus Sentence
Beginning of Overt Recall
(End of Trial)
End of Overt Recall
Two-second inter-trial interval
(Beginning of Next Trial)
Latency Interval Measure: (reflects, in part, operation of noun-phrase segmentation strategy)
Error Recall Measure: (reflects, in part, operation of different strategies from those used in comprehension)
TIME: